

PCT

世界知的所有権機関  
国際事務局  
特許協力条約に基づいて公開された国際出願

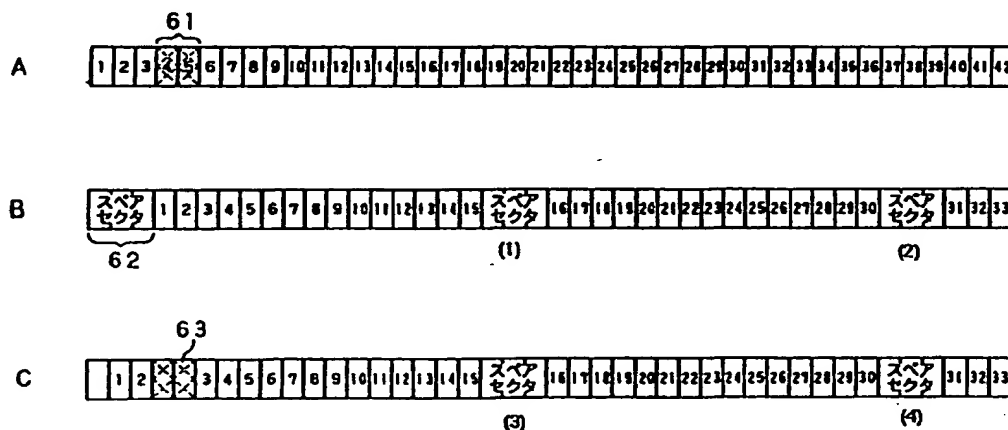
Reference



(51) 国際特許分類 G11B 20/12, 20/10	A1	(11) 国際公開番号 <b>WO98/36414</b>  (43) 国際公開日 1998年8月20日(20.08.98)
(21) 国際出願番号 PCT/JP98/00627  (22) 国際出願日 1998年2月16日(16.02.98)  (30) 優先権データ 特願平9/30630      1997年2月14日(14.02.97)      JP  (71) 出願人 (米国を除くすべての指定国について) ソニー株式会社(SONY CORPORATION)[JP/JP] 〒141-0001 東京都品川区北品川6丁目7番35号 Tokyo, (JP) (72) 発明者 ; および (75) 発明者 / 出願人 (米国についてのみ) 前田 哲(MAEDA, Satoshi)[JP/JP] 〒141-0001 東京都品川区北品川6丁目7番35号 ソニー株式会社内 Tokyo, (JP) (74) 代理人 弁理士 小池 晃, 外(KOIKE, Akira et al.) 〒105-0001 東京都港区虎ノ門二丁目6番4号 第11森ビル Tokyo, (JP)		(81) 指定国      JP, US.  添付公開書類 国際調査報告書

(54) Title: DATA RECORDING/REPRODUCING METHOD AND DEVICE

(54) 発明の名称 データ記録再生方法及び装置



62, (1), (2), (3), (4) ... spare sectors

## (57) Abstract

Each of the tracks of each disc in a plurality of disc drives is divided into blocks and spare sectors are provided between the blocks at predetermined intervals. Blocks including defective sectors are rearranged in a substitute region designated in accordance with a defect list in which the addresses of the detected defective sectors are registered and the preset correspondence relations in such a way that the sectors except the defective sectors are continuous on the disc. When the rearranged data are reproduced, all the data within a range designated to be read are consecutively read out and stored in a memory. When data are read out of the memory, the data of the defective sectors are excluded. When data are recorded, the data are consecutively recorded in a range designated to be written and dummy data are written in the defective sectors. Thus, the defects can be coped with while all the disc drives are not stopped and, further, the operating speed is not lowered after the defects are excluded.

(57) 要約

複数のディスクドライブの各ディスクのトラックを複数のブロックに分割し、ブロック間にスペアセクタを所定の間隔で配置する。検出された欠陥セクタのアドレスを登録する欠陥リストと、予め設定された対応関係とに基づいて指定される代替領域に、欠陥セクタを含むブロックを、セクタが欠陥セクタを除いてディスク上で連続して配置されるように置き換える。置き換えられたデータの再生時には、読出し指定された範囲内の全データを連続的に読出してメモリに格納し、メモリからデータを読み出す際に欠陥セクタのデータ削除する。また、記録時には、書込み指定された範囲内にデータを連続的に書込み、このとき欠陥セクタにはダミーデータを書き込む。これによって、全てのディスクドライブを停止することなく欠陥処理ができ、しかも欠陥処理後に動作速度が低下しないデータ記録再生方法および装置を提供することができる。

PCTに基づいて公開される国際出願のパンフレット第一頁に掲載されたPCT加盟国を同定するために使用されるコード (参考情報)

AL	アルバニア	FR	フランス	LT	リトアニア	SN	セネガル
AM	アルメニア	GB	イギリス	LV	ラトヴィア	SD	スーダン
AT	オーストリア	GE	ジョージア	MC	モナコ	TG	トーゴ
AZ	アゼルバイジャン	GH	ガーナ	MD	マルダヴィア	TJ	タジキスタン
BB	バハマ	GN	ギニア	MG	マダガスカル	TM	トルクメニスタン
BF	ブルキナファソ	GW	ギニアビサウ	MK	マケドニア共和国	TR	トルコ
BG	ブルガリア	GR	ギリシャ	ML	マリ	TT	トリニダード・トバゴ
BJ	ベナン	HU	ハンガリー	MN	モンゴル	TU	トルクメニスタン
BR	ブラジル	ID	インドネシア	MR	モーリタニア	UG	ウガンダ
BY	ベラルーシ	IE	アイルランド	MX	メキシコ	US	米国
CC	中央アフリカ共和国	IL	イスラエル	NE	ニジェール	UY	ウルグアイ
CG	コンゴ共和国	IN	インド	NL	オランダ	VN	ベトナム
CH	スイス	IT	イタリア	NO	ノルウェー	ZW	ジンバブエ
CI	コートジボワール	JP	日本	NZ	ニュージーランド		
CM	カメルーン	KE	ケニア	PL	ポーランド		
CN	中国	KR	韓国	PT	ポルトガル		
CO	コロンビア	KZ	カザフスタン	PR	プエルトリコ		
CR	コスタリカ	KG	キルギス	RO	ルーマニア		
DE	ドイツ	KK	カザフスタン	RS	セルビア		
EE	エストニア	KL	クウェート	SC	スコットランド		
		LL	リベリア	SE	スウェーデン		
		LS	レソト	SI	スロベニア		
				SK	スロバキア		
				SL	シエラレオネ		

## 明 細 書

### データ記録再生方法及び装置

#### 技 術 分 野

本発明は、デジタルデータを複数のディスク装置に記録し再生するデータ記録再生方法及び装置に関するものである。

#### 背 景 技 術

近年、CATV（ケーブルテレビ）等の普及による情報提供の多チャンネル化に伴い、従来のVTR（ビデオテープレコーダ）とは異なり、1台の映像・音声データ記録再生装置から複数の映像・音声データを同時に記録したり、再生したり、さらには記録しながら再生したり等の要求が高まりつつある。そして、この要求を満たすためにハードディスク等のランダムアクセス可能な記録媒体を用いて映像・音声を記録再生するビデオサーバ（または、AV（Audio and/or Video）サーバとも呼ばれる）と呼ばれる装置が普及しつつある。

一般的に、放送局内におけるビデオサーバは、画質や音質に対する要求から、必要とされるデータの転送レートが高い上に長時間のデータを記録するために大容量である必要がある。そこで、映像・音声データを蓄積するとともに並列処理が可能な複数のハードディスク（以下HD）装置を含むデータ記録再生装置を用いることによ

りデータの転送レートの高速化と大容量化を図る試みや、更にパリティデータを記録しておくことにより、万一いずれかのH D 装置が故障しても信頼性を確保できるようにする試みがなされている。これにより、放送局が提供しようとしている番組の内容や放送形態により要求されているチャンネル数が異なる場合であっても、複数の音声・映像データからなる素材データを分散的に記録しておき多チャンネル送出を同時に行ったり、同一の素材データを再生時間をずらして多チャンネルで再生することにより、V O D (ビデオオンデマンド)やN V O D (ニアビデオオンデマンド)等のシステムを構築する等、多様な使用形態に対応することができるマルチチャンネルビデオサーバを実現することができる。

このようなビデオサーバに用いられるデータ記録再生装置には、1988年 Patterson等によって発表された論文("A Case for Redundant Arrays of Inexpensive Disks (RAID)", ACM SIGMOND Conference, Chicago, Ill, Jun. 1-3, 1988)に提唱されている複数のH D からなるハードディスクドライブ(以下H D D と記す)をさらに複数台用いて構成されたR A I D (Redundant Arrays of Inexpensive Disks) 技術が用いられている。

上記論文の中でR A I D は、R A I D - 1 からR A I D - 5 まで5つに分類されている。R A I D - 1 は2つのH D D に同じ内容を書込む方式である。R A I D - 3 は、入力データを一定の長さに分割して複数のH D D に記録するとともに、各H D D の互いに対応するデータブロックの排他的論理和であるパリティデータを生成して他の1台のH D D に書込む方式である。さらにR A I D - 5 は、データの分割単位(ブロック)を大きくして、1つの分割データをデ

ータブロックとして1つのHDDに記録するとともに、各HDDの互いに対応するデータブロックの排他的論理和をとった結果（パリティデータ）をパリティブロックとして他のHDDに記録するとともにパリティブロックを他のHDDに分散する方式である。

図1は、RAIDの基本的な構成を示すブロック図である。このRAID10は、SCSIプロトコルコントローラ（以下SPCと記す）1、ストライピング部2、ユニフィケーション部3、メモリ部4、SPC部5、HDD部6、CPU7を備えて構成されている。ここでは、HDD部6がHDD<sub>1</sub>～HDD<sub>8</sub>までの8台のHDDから構成され、データの入出力インタフェースにはSCSI（Small Computer System Interface）を採用している場合を例として、以下に各部の構成および機能を説明する。

SPC1は、RAID10のデータの入出力を管理・制御するための入出力制御手段であり、その動作はCPU7により制御されている。またSPC1は、SCSIを介してイニシエータ8に接続されている。

ストライピング部2は、SPC1を介して入力された記録データを、HDD部6のHDD<sub>1</sub>～HDD<sub>8</sub>に分配して記録するためのデータ分配手段である。各HDDに分配されるデータには、パリティデータが付加される。

ユニフィケーション部3は、HDD部6のHDD<sub>1</sub>～HDD<sub>8</sub>から読み出されたデータを1つの再生データに構成するためのデータ構成手段である。

メモリ部4は、HDD部6の記録／再生データを一時的に記憶するためのバッファメモリであり、半導体メモリ素子を用いて構成さ

れる。

SPC部5は、HDD部6とメモリ部4との間のデータ転送を制御する転送制御手段であり、SPC1と同様にCPU7により制御されている。

HDD部6は、上述のようにHDD<sub>1</sub>からHDD<sub>8</sub>までの8台のHDDから構成されている。これらのHDDには、通常、3.5インチ径や5.25インチ径程度の小径ディスクが用いられる。各HDDには、後述するように、同一のアドレスが割り当てられ、データが並列に記録および再生される。

メモリ部4とSPC部5は、ともに、HDD部6のHDD<sub>1</sub>～HDD<sub>8</sub>に対応するメモリ<sub>1</sub>～メモリ<sub>8</sub>およびSPC<sub>1</sub>～SPC<sub>8</sub>を備えている。そして、HDD部6の各HDDのデータ入出力は、メモリ部4およびSPC部5を介して行われる。

また、イニシエータ8は、SPC1とSCSIを介して接続され、各種制御コマンドを発行して、RAID10の動作を制御するものである。

RAID10へ入力データを記録する時には、SPC<sub>1</sub>を介して入力されたデータに対して、ストライピング部2でストライピング時にパリティ符号が付加され、メモリ部4、SPC部5を介してHDD部6の各HDDに分配される。また、RAID10からデータを再生する時には、記録する時とは逆の手順により、HDD部6の各HDDからデータが読み出され、ユニフィケーション部3で再生データが構成される。

HDD<sub>2</sub>～HDD<sub>8</sub>のHDにおいて、データはセクタと呼ばれる単位で均等に分配され記録される。すなわち、図1のRAID10で

は、HDD部6のHDD<sub>1</sub>～HDD<sub>8</sub>までの各セクタに、データが均等に分配される。またHDD部6は、上述したRAID-3またはRAID-5の構成をとっているため、HDD部6のHDD<sub>1</sub>～HDD<sub>8</sub>のうちの1台に何らかの障害が発生して正常なデータ転送ができなくなっても、残り7台のHDDに分配されて記録されているデータからパリティデータを用いて必要なデータを復元して、正常に転送することができる。これにより、RAID10は、1台のHDDに障害が発生した場合にも、機能を停止せずに運用を継続することができる。

しかし、HDD<sub>1</sub>～HDD<sub>8</sub>のうちの、2台以上のHDDの同一アドレスに障害が発生した場合には、必要なデータを復元して正常に転送することができなくなってしまうことがある。このため、1台のHDDに障害が発生したときには、障害により読み出せなくなったデータを、2台目のHDDに障害が発生する前に復元して、復元データを正常な代替セクタに記録し直す欠陥処理が行われる。この欠陥処理は、CPC7からの「リアサインブロック (Reassign Block) コマンド」により行われている。

図2は、リアサインブロックコマンドにより、欠陥処理が行われる様子を説明する図である。すなわち、欠陥が発生したセクタ21のアドレスを、予め用意された別の正常なスペアセクタ30に再登録（以下この処理をリアサイン; Reassignと記す）した後、欠陥セクタ21のデータ復元処理を行い、復元されたデータをスペアセクタ30に再記録（以下この処理をリビルト; Rebuilt と記す）するというものである。このスペアセクタ30は、同一ディスク上の別のトラックに設けられたり、別のディスク上に設けられたりする。

ところが、上記のリアサインブロックコマンドを実行すると、欠陥セクタ 21 のみがスベアセクタ 30 にリアサインされるため、RAID 10 のセクタアドレスのディスク上での物理的順番が変わってしまう。このため、データアクセス時に、欠陥処理前には必要なかった「シーク＋回転待ち」時間が発生してしまう。すなわち、ディスク上での物理的順番が変わってしまうので、データにアクセスするときスベアセクタ 30 にヘッドを移動させ（シーク）、さらに所望のスベアセクタ 30 がヘッド位置に回ってくるまで（回転待ち）の時間が発生するのである。従って、リアサインブロックコマンドの実行後には、データアクセス時間が著しく長くなってしまうことがある。転送されるデータが、リアルタイム性を要求されるビデオ／オーディオデータ等である場合には、所定の時間内に必要なデータを転送することができなくなってしまう、再生画像／音声が乱れる原因になる。

また、複数の HDD を用いて構成される RAID では、各 HDD がデータを並列に転送することにより、1 チャンネル分の画像または音声データを転送するために必要な転送速度よりも高速でデータを転送することが可能である。例えば、4 倍速の高速転送が可能な場合には、1 台の RAID が 4 つのデータ転送を並列に行うことが可能になり、ユーザ A からユーザ D までの 4 ユーザが、同時にアクセスすることができる。

図 3 は、RAID 10 において 4 倍速のデータ転送を行っている様子を説明している。すなわち、ディスク上に記録された 1-1～1-2, 2-1～2-2, 3-1～3-2, 4-1～4-2 の 4 種類のデータを各々 4 倍速で読み出し、メモリ部 4 に格納する。そして、このメモリ部



4 から、上記 4 種類のデータをそれぞれ 1 倍速で読み出す。このようにすれば、ユーザ A からユーザ D までの 4 ユーザが、同時に異なるデータにアクセスすることができる。

しかし、上記のように構成された R A I D において、前述のリアサインブロックコマンドを実行する場合には、R A I D 全体の動作を一旦止めて、欠陥処理を行う H D D を切り離さなければならない。これは、リアサインブロックコマンドが、H D D 単体にアクセスするコマンドであるためである。図 3 のように 1 台の R A I D 1 0 を 4 ユーザが使用している場合には、4 ユーザが全ての作業を止めなければならないため、R A I D 運用上の大きな制限事項となっている。

## 発 明 の 開 示

本発明は、上述した問題点に鑑みてなされたものであり、運用を停止することなく欠陥処理を行うことができ、しかも欠陥処理後にディスクアクセスが遅くならない R A I D のデータ記録再生方法及び装置を提供することを目的とする。

本発明に係るデータ記録再生方法は、複数のディスクから構成されたディスクドライブの各ディスクのトラックが複数のブロックに分割され、上記ブロックの間には所定のセクタ数からなるスペアセクタが所定の間隔で予め配置され、上記ディスクドライブにデータを分配して記録及び／又は再生するデータ記録再生方法であって、上記ディスクドライブの欠陥セクタを検出する欠陥セクタ検出工程と、検出された欠陥セクタのアドレスを欠陥リストに登録する欠陥

セクタ登録工程と、上記欠陥セクタを含むブロックを上記欠陥リストと予め設定された対応関係とに基づいて、欠陥セクタを除いてディスク上で連続して配置されるよう上記スペアセクタを用いて代替領域に再登録するリアサイン工程とを有することを特徴としている。

上記再登録されたブロックに含まれている欠陥セクタのデータを復元する復元工程と、上記復元データを上記代替領域に再登録されたブロックに再記録するリビルト工程とをさらに有することが挙げられる。

上記リアサイン工程では、上記欠陥セクタ数が上記スペアセクタ数よりも多いときには、上記欠陥セクタを含むブロックの再登録を複数のスペアセクタを用いて行うことが挙げられる。

また、本発明に係るデータ記録再生装置は、複数枚のディスクから構成されたディスクドライブの各ディスクのトラックが複数のブロックに分割され、上記ブロックの間には所定のセクタ数からなるスペアセクタが所定の間隔で予め配置され、上記ディスクドライブにデータを分配して記録及び／又は再生するデータ記録再生装置において、上記ディスクドライブの論理セクタと物理セクタとの対応を記憶するアドレスマッピング手段と、上記ディスクドライブの欠陥セクタを検出して、該欠陥セクタのアドレスを記憶する欠陥セクタ記憶手段と、上記欠陥セクタを含むブロックを上記アドレスマッピング手段と上記欠陥セクタ記憶手段とに基づいて、上記ディスク上で上記欠陥セクタを除いて連続して配置されるように上記スペアセクタを用いて代替領域に再登録するリアサイン手段とを有することを特徴としている。

上記再登録されたブロックに含まれている欠陥セクタのデータを

復元する復元手段と、上記復元されたデータを、上記代替領域に再登録されたブロックに再記録するリビルト手段とをさらに有することが挙げられる。

上記複数のディスクドライブの記録／再生データを記憶する記憶手段をさらに備え、上記代替領域に再記録された復元データを再生するときには、読出し指定された範囲内の、スベアセクタ及び欠陥セクタを含む全セクタからデータを連続的に読出して上記記憶手段に格納し、格納されたデータを読み出すときに上記欠陥セクタのデータを削除することが挙げられる。また、上記欠陥セクタを含む領域にデータを記録するときには、書込み指定された範囲内の、スベアセクタ及び欠陥セクタを含む全セクタにデータを連続的に書込み、上記欠陥セクタにはダミーデータを書き込むことが挙げられる。

これによって、欠陥処理後のヘッドのシーク動作及び回転待ち動作によるディスクへのアクセス劣化が発生せず、しかも R A I D が通常動作を行いながら欠陥処理の割り込みが可能となる。すなわち、マルチチャンネルの使用においては、そのうちの 1 チャンネル分だけを使って欠陥処理を行うことができ、残りのチャンネルの運用を止める必要がなくなる。

さらに、1 チャンネル分を欠陥処理専用として使用可能な場合には、イニシエータが自動で欠陥処理コマンドを、この専用チャンネルに発行するように制御すれば、欠陥処理が自動で実施され、ユーザは全く欠陥を意識することなく運用継続できる。

図面の簡単な説明

図1は、RAIDの基本的な構成を示すブロック図である。

図2は、リアサインブロックコマンドにより、欠陥処理が行われる様子を説明するための図である。

図3は、RAID10において4倍速のデータ転送が行われる様子を説明するための図である。

図4は、RAIDを構成するHDDの基本的なセクタ構造の一例を示す図である。

図5A、Bは、RAID10のHDDに対するセクタアドレス割当の一例を示す図である。

図6A～Dは、RAID論理アドレスマッピング、HDD欠陥リスト、スペアアドレスリスト、RAID欠陥アドレスリストの構成を示す図である。

図7A～Cは、欠陥セクタ数が1つのスペアセクタ数よりも小さい場合の置換え処理について説明するための図である。

図8A～Eは、上記の置換え処理を行う際に参照される、RAID論理アドレスマッピング、HDD欠陥リスト、HDDスペアリスト、RAID欠陥リストである。

図9は、1ブロックのスペアセクタを用いる置換え処理の手順を示すフローチャートである。

図10A～Cは、複数ブロックのスペアセクタを用いる置換え処理について説明するための図である。

図11A～Eは、上記の置換え処理を行う際に参照される、RAID論理アドレスマッピング、HDD欠陥リスト、HDDスペアリスト、RAID欠陥リストである。

図12A～Cは、データ読出し時のスキップ動作について説明す

るための図である。

図13は、データ読出し時の置換え処理およびスキップ動作の手順を示すフローチャートである。

図14A～Cは、データ書込み時のダミーデータ付加処理動作について説明するための図である。

図15は、データ書込み時のダミーデータ付加処理の手順を示すフローチャートである。

### 発明を実施するための最良の形態

以下、本発明に係るデータ記録再生方法及び装置の好ましい実施の形態について、図面を参照しながら説明する。

ここで、本発明の具体的な説明に先だて、HDDの基本的なセクタ構造について先ず説明する。

図4は、RAIDを構成するHDDの基本的なセクタ構造の一例を示している。このHDDは、計10枚のHDを備え、各HDの表側および裏側にアクセスする20本のヘッド（ヘッド0からヘッド19）を備えている。

また、1枚のディスクは、ディスクの外周から内周に向けてゾーン1からゾーン7までの7つのゾーンに分かれており、したがって内周側のゾーンほどセクタが少なくなっている。

同じゾーンに属するトラックはシリンダと呼ばれ、各ゾーンは複数のシリンダからなる。すなわち、ゾーン1は1177個のトラック（シリンダ）からなり、ゾーン2は573個のトラック、・・・となる。したがって、1枚のディスクには、合計5272（シリン

ダ) が設けられている。そして、各シリンダにはヘッド 0 ～ 19 までのヘッドがアクセスし、各ヘッドは各トラックにアクセスする。

次に、本発明のデータ記録再生方法の特徴である、欠陥処理について図 1 を用いて説明する。イニシエータ 8 から、RAID 10 のセクタアドレスに対して、「データ読出し (READ) コマンド」や「書込み (WRITE) コマンド」が発行されると、RAID 10 は、そのセクタを構成している HDD 部 6 の各 HDD のセクタアドレスに対して、これらのコマンドを実行する。通常は、RAID 10 のアドレス番号と HDD 部 6 の各 HDD のアドレス番号とは同一にされており、CPU 7 には、後述する RAID 論理アドレスマッピング表および各 HDD の欠陥アドレスリストと、それに対応した RAID 10 の欠陥アドレスリストが設けられ、これらのリストを用いて欠陥処理が行われる。

すなわち、本発明に係る欠陥処理は、欠陥セクタを代替領域に 1 対 1 にマッピングするのではなく、上述したマッピング表およびリストを参照することにより、数セクタずつ所定の間隔に配置された配置されたスベアを用いてマッピングを行うことを特徴としている。以下では、このマッピングを中心に説明する。

図 5 A, B は、RAID 10 の HDD へのセクタアドレスへの割当の一例を示している。

図 5 A は、HDD の 1 トラック分のセクタアドレスを示している。この例では 1 トラックが 42 セクタで構成されている。

図 5 B は、上記の HDD セクタアドレスに対する RAID アドレスの割当を示している。RAID のセクタアドレス割当と、HDD のセクタアドレスとは、図 1 に示すようにイニシエータ 8 から RA

RAID 10 へのコマンドに例えば RAID のセクタアドレスに相当する情報が含まれて入力されると、RAID 10 の CPU は HDD へのセクタアドレスに変換して、各 HDD に対するコマンドを実行させるようになされている。図 5 B に示す RAID のセクタアドレスは、RAID 10 の外部からの RAID に対するセクタアドレスであり、図 5 A に示す HDD のセクタアドレスは RAID 内から見た各 HDD のアドレスを示すのである。図 5 B に示すように RAID のセクタアドレス割当は、ある連続した複数のセクタを 1 ブロックとして、そのブロック間にスペアセクタとして数セクタを所定の間隔で配置するというものである。この例では、1 トラックを 42 セクタ、1 ブロックを 15 セクタ、ブロック間のスペアセクタを 3 セクタ、トラック間のスペアセクタを 6 セクタとした場合のセクタ割当を示している。なお、1 ブロックのセクタ数およびスペアセクタ数は、パラメータにより設定可能である。

上述した例は、説明のため 1 ブロック 15 セクタで、ブロック間に 3 セクタのスペアセクタを設けた例を示したが、実際には、全データに対して約 0.4 ~ 1 % のスペアセクタを準備することになるので、300 ~ 500 セクタ毎に 3 ~ 5 セクタのスペアセクタを配置する。すなわち、欠陥セクタに対するデータ復元処理を行い、復元されたデータをスペアセクタに再記録する。ここで、ブロックを大きくし過ぎると、データをリビルトするときにはデータを蓄積するためのメモリ容量が大きくなり過ぎるため好ましくない。512 K バイト (Byte) 程度のメモリ容量でカバーできるようなブロックサイズが適当である。

そして、RAID 10 の、あるセクタに欠陥が発生したときには、

前述したRAID論理アドレスマッピング表および欠陥リストを参照して、RAID10のアドレスの物理的順番が変わらないように置換え処理を実行する。この置換え処理は、図2に示したような、欠陥セクタだけをスペアセクタに置き換える従来の欠陥処理とは異なり、ブロック間に存在するスペアセクタやブロック内に発生した欠陥セクタを、うまくスキップするように行われる。これにより、従来の欠陥処理後に発生していた、「シーク+回転待ち」時間の追加によるデータアクセスの遅れを防ぐことができる。このスキップ処理については後述する。

また、上記の置換え処理は、RAID10のコマンド一つとして実行処理できるので、欠陥が発生したHDD単体にアクセスするために、欠陥処理を行うHDDをRAID10から切り離す必要がない。すなわち、外部からイニシエータを介してRAID10へのコマンド1つで、RAID10内ではCPU7が上述したマッピング表及びリストを参照して、置換え処理を各HDDに対して行うために欠陥処理を行うHDDがRAID10から切り離されない。従って、RAID10全体の動作を停止することなしに欠陥処理を行うことができる。

このため、図3のように、1台のRAIDを4ユーザが使用する場合には、1ユーザ分のチャネルだけを使用して欠陥処理を行い、残りの3ユーザは運用を続けることができる。さらに、1ユーザ分のチャネルを欠陥処理専用として使用できる場合には、イニシエータ8が、欠陥発生時に自動的に欠陥処理コマンドをこの専用チャネルに対して発行するように制御すれば、ユーザが運用中に欠陥が自動的に処理される。これにより、ユーザは、欠陥の発生を意識する



ことなく R A I D の運用を継続できる。

図 6 A ～ D は、上述した、R A I D 論理アドレスマッピングおよび H D D の欠陥リストおよびスペアアドレスリストと、それに対応した R A I D 1 0 の欠陥アドレスリストを示している。これらは、前述したように C P U 7 に格納される。実際には、C P U 7 は、中央演算処理装置と、上述したリスト等が格納された R O M と演算のためにデータが格納される R A M から構成されている。本実施例では説明を簡単にするため中央処理装置、R O M、R A M をまとめて C P U 7 とし、以下説明する。欠陥アドレス発生時は、この R A I D 論理アドレスマッピングおよび欠陥リストを参照して、R A I D 1 0 のアドレスの物理的順番が崩れないよう（変わらないよう）にセクタの置換え処理が行われる。

図 6 A は、R A I D 論理アドレスマッピング表を示している。このマッピング表は、R A I D 1 0 の論理セクタアドレスと、R A I D 1 0 を構成する H D D 部 6 の各 H D D 単体のセクタアドレスとの対応関係を示すためのものである。

図 6 B は、H D D 欠陥リストを示している。このリストは、何らかの障害が派生した R A I D を構成する各 H D D のセクタアドレスを登録するためのものである。

図 6 C は、H D D スペアリストを示している。このリストは、R A I D に割り当てられたセクタアドレスで、スペアセクタに割り当てられたアドレスに対応する各 H D D のセクタアドレスを示している。例えば、図 5 B でスペアセクタとして割り当てられたアドレスに対応する H D D のセクタアドレスは、図 5 A では 1, 2, 3 及び 1 9, 2 0, 2 1 であるので、このアドレス値が図 6 C に示すスペ

アリストに登録される。このアドレス値は全ての処理が行われる前に登録されていても良いし、動作中上位からの指示により変更することもできる。

図6Dは、RAID欠陥リストを示している。このリストは、RAID10としての欠陥を登録するためのものである。すなわち、欠陥が生じたHDDのセクタアドレスに対応するRAIDのセクタアドレスを登録するためのリストであり、例えば図5Aのアドレス4, 5に欠陥が生じたとき、対応するRAIDのアドレスは1, 2であるのでこの値が登録される。登録された欠陥セクタは、置換え処理後に削除される。

上記のセクタの置き換えによる欠陥処理を行う場合には、ブロック間に配置されるスペアセクタや、ブロック内に発生した欠陥セクタをうまくスキップするように置換え処理を行わないと、前述の「シーク+回転待ち」時間が発生してしまい、データアクセスの遅れを防ぐことができなくなってしまう。以下では、この置換え処理の手順について説明する。

図7Aに示すように、HDD<sub>1</sub>のセクタアドレス4及び5に欠陥が発生した場合を例として考える。このとき、RAID10に割り当てられたセクタアドレスは、図7Bのように割り当てられているとすると、図7Cのように、HDD<sub>1</sub>に対してRAID10のアドレスとして割り当てられた1, 2セクタを、トラックの先頭に配置された3セクタ分のスペアセクタ62にシフトさせ、欠陥が発生した2セクタ分の領域61を空領域63にする。

このように、欠陥セクタ数が1つのスペアセクタ数よりも小さい場合には、同一トラック内でセクタをスペアセクタにシフトさせる

ことにより置換え処理を行うことができる。この置換え処理によれば、欠陥処理後にも、ディスク上のトラック内におけるデータの物理的順番が変わらないため、従来の欠陥処理の問題であった、スベアセクタにアクセスするための余分なシーク時間が発生しない。なお、シフトさせたセクタに対するデータの読み出し及び書き込みに関する詳細は後述する。

図8A～Eは、上記の置換え処理を行う際に参照される、RAID論理アドレスマッピング表、HDD欠陥リスト、HDDスベアリスト、RAID欠陥リストを示している。

図8Aは置換え処理前のRAID論理アドレスマッピングであり、図8Bは置換え処理後の、セクタ割当が変更されたRAID論理アドレスマッピングである。置換え処理前のマッピングの領域71には、それぞれRAIDセクタアドレスが1および2に対応するHDD<sub>1</sub>のセクタアドレスは4および5が登録されている。置換え処理後のマッピングの領域72には、RAIDセクタアドレスが1および2に対応するHDD<sub>1</sub>のセクタアドレスがそれぞれ2および3に置き換えられている。すなわち、上述した置き換え処理を行うことでRAID論理アドレスマッピング表の内容も処理に応じて書き換えられるのである。

図8Cは、HDD欠陥リストであり、欠陥が発生したHDD<sub>1</sub>のセクタアドレス4、5が領域73に登録される。

図8Dは、HDDスベアリストであり、置換え処理に用いられたHDD<sub>1</sub>の2セクタ分のスベアセクタが領域74から削除されている。

図8Eは、RAID欠陥リストであり、欠陥が発生したRAID

セクタである、セクタ 1 とセクタ 2 が登録されている。登録された欠陥セクタは、割当変更後に削除される。

次に、上述した、1 ブロックのスベアセクタを用いる置換え処理の手順を説明する。

図 9 は、この置換え処理の手順を示すフローチャートである。

まず、セクタの置き換え処理が開始されると、処理はステップ S T 1 に移る。ステップ S T 1 では、欠陥が発生した H D D のセクタマトレスが C P U 7 に通知される。C P U 7 への通知は、C P U 7 から S P C を経由して各 H D D にコマンドを送り、このコマンドに対する応答として通知されてもよいし、各 H D D が定期的に C P U 7 に送られる。データに基づいて通知されて欠陥発生が C P U に通知される。図 7 A ~ C の例では、H D D<sub>1</sub> のセクタアドレス 4, 5 が C P U 7 に通知される。次いで処理はステップ S T 2 に移る。

ステップ S T 2 では、H D D 欠陥リストにステップ S T 1 で C P U 7 に通知されたセクタアドレスを登録する。図 7 A ~ C の例では、H D D<sub>1</sub> のセクタアドレス 4, 5 が登録され、したがって図 8 C に示すリストが作成される。次いで処理はステップ S T 3 に移る。

ステップ S T 3 では、R A I D 論理アドレスマッピング表を参照して、ステップ S T 2 で H D D 欠陥リストに登録されたセクタアドレスに対応する R A I D のセクタアドレスを検索する。次いでステップ S T 4 にて、R A I D 欠陥リストに検索結果を登録する。図 7 A ~ C の例では、図 8 C に示すリストから図 8 A を検索して R A I D のセクタアドレスおよび 2 を得て、これを図 8 E の R A I D 欠陥リストに登録する。

次いで処理はステップ S T 5 に移り、ステップ S T 5 では R A I

D アドレスマッピング表を上記した欠陥リストに基づいて書き換える。

上記の例では、図 8 A の枠で囲んだ部分 7 1 を図 8 B の枠で囲んだ部分 7 2 に書き換えて図 8 B が作成される。

次いで処理はステップ S T 6 に移り、リビルト処理を行う。すなわち、H D D<sub>1</sub> のセクタアドレス 4, 5 のデータを復元して、復元データを H D D<sub>1</sub> のセクタアドレス 2, 3 に書き込む。次いで処理はステップ S T 7 に移る。

ステップ S T 7 では、R A I D 欠陥リストに登録されているマトレスを削除し、H D D スペアリストに登録されているアドレスがマッピング表で書き換えたアドレスを消去する。上記の例では、図 8 E に示す R A I D 欠陥リストからセクタ 1, 2 を削除し、図 8 B に示すように書き換えたマッピング表のアドレスを図 8 D から削除、すなわち H D D<sub>1</sub> のアドレス 2, 3 を書き換えたので、図 8 D に示すスペアリストの H D D<sub>1</sub> のアドレス 2, 3 を削除する。

以上の手順により、1 ブロックのスペアセクタを用いる置換え処理が終了する。

上記した、1 ブロックのスペアセクタを用いる置換え処理では、欠陥セクタ数がスペアセクタ数を越えた場合には、欠陥処理ができなくなってしまう。そこで、このような場合には、複数ブロックのスペアセクタを用いて、同様の置換え処理を行う。以下に、複数ブロックのスペアセクタを用いる置換え処理について具体的に説明する。

図 1 0 A に示すように、H D D<sub>1</sub> のセクタマトレスが 4 から 1 0 までの 7 セクタに亘って欠陥が発生した場合を例として考える。こ

のとき、RAID 10のアドレスは、図10Bのように割り当てられている。この割当は、1ブロックのスベアセクタを用いる置換え処理を行う場合と同様である。

トラックの先頭に配置された3セクタ分のスベアセクタ82に、RAIDのセクタアドレス1から3までをシフトさせ、欠陥セクタ4～6までの3セクタ分を空領域にする。また、各々3セクタ分のスベアセクタ83，84を用いて、4～30セクタをシフトさせ、欠陥セクタ7～10までの4セクタ分を空領域にする。これにより、上記の欠陥が発生した7セクタ分の領域81を、空領域85にして、RAIDのセクタアドレスの割当を変更する。

このように、欠陥セクタ数が1つのスベアセクタ数よりも大きい場合には、複数のスベアセクタを用いてセクタデータを同一トラック内でシフトさせることにより置換え処理を行うことができる。この置換え処理によれば、欠陥処理後にも、ディスク上のトラック内におけるデータの物理的順番が変わらないため、従来の欠陥処理の問題であった、スベアセクタにアクセスするための余分なシーク時間が発生しない。

図11A～Eは、上記の置換え処理を行う際に参照される、RAID論理アドレスマッピング、HDD欠陥リスト、HDDスベアリスト、RAID欠陥リストを示している。これらは、前述したように、RAID 10のCPU 7内に構成されるものであり、前述の、1ブロックのスベアセクタを用いる置換え処理を行う際に参照されるものと同様である。

図11Aは置換え処理前のRAID論理アドレスマッピングであり、図11Bは置換え処理後の、割当が変更されたRAID論理ア

ドレスマッピングである。置換え処理後のマッピングの領域 9 2 には、欠陥が発生した H D D<sub>1</sub> の欠陥セクタである 4 ~ 6 セクタが 1 ~ 3 セクタに置き換えられ、欠陥セクタ 7 ~ 1 0 セクタが 1 1 ~ 1 4 セクタに置き換えられることを示している。これは、図 1 0 B の R A I D アドレスの 1 ~ 3 セクタに対する H D D<sub>1</sub> のセクタ割当が、図 1 0 A のセクタ 4 ~ 6 から、図 1 0 C に示すように、元のセクタ 1 ~ 3 に変更されることを示している。同様に、図 1 0 B の R A I D アドレスの 4 ~ 7 セクタに対する H D D<sub>1</sub> のセクタ割当が、図 1 0 A のセクタ 7 ~ 1 0 から、図 1 0 C に示すように、元のセクタ 1 1 ~ 1 4 に変更されることを示している。

図 1 1 C は、H D D 欠陥リストであり、欠陥が発生した H D D<sub>1</sub> の 4 ~ 1 0 セクタが領域 9 3 に登録されている。

図 1 1 D は、H D D スペアリストであり、置換え処理に用いられた H D D<sub>1</sub> の 7 セクタ分のスペアセクタである、1, 2, 3, 1 9, 2 0, 2 1, 3 7 セクタが領域 9 4 から削除されている。

図 1 1 E は、R A I D 欠陥リストであり、欠陥が発生した R A I D セクタアドレス 1 ~ 7 が登録されている。登録された欠陥セクタは、割当変更後に削除される。

次に、上述した、複数ブロックのスペアセクタを用いる置換え処理の手順を説明するが、置き換え処理の手順そのものは図 9 の単一ブロックのスペアセクタを用いた置き換え処理の手順と同じである。

上述した H D D<sub>1</sub> の 4 ~ 1 0 までの 7 つのセクタに欠陥が発生した場合の置換え処理を行った場合について、再び図 9 を用いて説明することにする。

まず、置き換え処理が開始されると、処理はステップ S T 1 に移

り、欠陥が発生したHDDのセクタアドレスがCPUに通知される。上述の例では、欠陥が発生したHDD<sub>1</sub>のセクタアドレス4～10ステップST11では、HDD<sub>1</sub>から4、5、6、7セクタに欠陥が発生したことがRAID10に通知される。

次いで処理は、ステップST2に移り、CPU7に通知されたセクタアドレスをHDD欠陥リストに登録する。上述した図10A～Cの例では、HDD<sub>1</sub>のセクタアドレス4～10までがHDD欠陥リストに登録され図11Cに示すリストが作成される。

次いで処理はステップST3に移り、RAID論理アドレスマッピング表を参照して対応するRAIDのセクタアドレスを検索し、ステップST4で検索した結果をRAID欠陥リストに登録する。図10A～Cの例では図11Cから対応するRAIDセクタアドレスを図11Aのマッピング表で検索し、結果であるアドレス1から7を図11Eに登録する。

次いで処理はステップST5に移り、2つの欠陥リストに基づいて表を書き換える。上述の例では、RAIDアドレス1が30までHDD<sub>1</sub>のアドレスを書き換える。

次いで処理はステップST6に移って、リビルト処理を行ってHDD<sub>1</sub>のアドレス4～36までの復元データをHDD<sub>1</sub>のセクタアドレス8～40までに書き込む。

次いで処理はステップST7に移り、RAID欠陥リストからアドレス1～7を削除して、HDDスペアリストの1、2、3、19、20、21、37を削除する。

以上の手順により、複数ブロックのスペアセクタを用いる置換え処理が終了する。



このように、欠陥セクタの数に応じて、使用するスペアブロックの数を変更することにより、どのような欠陥にも、同一トラック内のセクタ置換え処理により対応することができる。しかし、欠陥の数が、トラック間に配置された複数のスペアセクタを用いても処理できない程多数の場合には、一般に、欠陥が加速度的に増加する傾向を示す。この場合には、欠陥処理により対応することは適当でなく、ディスクを交換することが必要である。

ディスクを交換したときには、交換ディスク全面のリビルトによりデータを復帰させるようにする。また、ブロックサイズ、およびブロック間のスペアセクタ数は、セクタ割当時にパラメータで設定可能であり、欠陥発生率等の状況に応じて適当な値に変更すればよい。

次に、上述した置き換え処理を行うときのデータの読み出しおよびデータ書き込み動作について説明する。

まず、データを読み出す時の動作について説明する。

図 1 2 A に示すように、HDD<sub>1</sub> のセクタアドレス 4, 5 に欠陥が発生した場合を再び例として考える。このとき、RAID 10 のアドレスは、図 1 2 B のように割り当てられている。ここでは、トラックの先頭およびブロック間に配置されるスペアセクタ 1 0 2, 1 0 3 に加えて、トラックの終端にもスペアセクタ 1 0 4 が配置されている場合を示している。このときには、図 1 2 C のように、1, 2 セクタを、トラックの先頭に配置された 3 セクタ分のスペアセクタ 6 2 にシフトさせ、欠陥が発生した 2 セクタ分の領域 1 0 5 を空領域にする置換え処理が行われる。

このため、図 1 2 B の RAID アドレスのうちの区間 A で示され

る範囲のセクタ 1～20 のデータを読み出そうとする場合には、HDD<sub>1</sub> のディスク上では、図 12C の区間 A' で示される空領域 105 を含む範囲が実際に読み出されることになる。このとき、セクタ 1, 2 を読み出した後に、空領域 105 をスキップしようとする、セクタ 3 の読出し開始位置を検出するためにディスクの「回転待ち」時間が必要になる。従って、データの転送に余分な時間が必要になってしまう。

具体的には、RAID 10 は、RAID 論理アドレスマッピングを参照して、HDD<sub>1</sub> のセクタ 2～3 と、セクタ 6～18 と、セクタ 22～26 を読み出し、また HDD<sub>2</sub> のセクタ 4～18 と、セクタ 22～26 を読み出そうとする。このとき、ブロック内の欠陥セクタおよびブロック間のスベアセクタを避けて読み出しを行うために、HDD<sub>1</sub> のセクタ 2～3, セクタ 6～8, セクタ 22～26 の 3 つの領域、および HDD<sub>2</sub> ～HDD<sub>7</sub> のセクタ 4～8, セクタ 22～26 の 2 つの領域に対して、“Read” コマンドを分けて発行してしまうと、コマンド間で必ず回転待ちが発生し、実行時間が遅れてしまう。

そこで、本発明のデータ記録再生方法では、区間 A' のセクタ 1 からセクタ 20 までの範囲を、空領域 105 およびスベアセクタ 106 も含めて連続して読み出して、読み出した全データをメモリ部 4 に一旦格納する。そして、メモリ部 4 に格納されたデータを読み出す際に、不要なデータである空領域 105 およびスベアセクタ 106 のデータを削除して出力する。

以上のスキップ処理により、置換え処理後の空領域を含む読出し範囲が読み出し指定された場合にも、HDD 単体の動作としては、

実際には欠陥セクタおよびスベアセクタのスキップ動作を行わないために、HDDの「回転待ち」時間が発生せず、最小限のデータアクセス時間でデータを読み出すことができる。

次に、上述した、スキップ処理によるデータの読出しをする時の手順について図13を用いて説明する。

図13は、このスキップ処理によるデータ読み出しの手順を示すフローチャートである。

この処理が開始されると、ステップST21に移る。ステップST21では、イニシエータ8から“Read”コマンドを受ける。上述した例ではRAIDのセクタアドレス1～20に対して“Read”コマンドが発行されCPU7がうける。次いで処理はスラップST22に移る。

ステップST22では、発行されたRAIDのセクタアドレスに対する、各HDDのセクタアドレスをRAID論理アドレスマッピング表を参照して検索し、次いでステップST23で検索した各HDDへの“Read”コマンドを発行する。図12A～Cの例では図8Aに示すRAID論理アドレスマッピングが参照され、RAID10の1～20セクタに対する読み出し(Read)コマンドが、HDD1ではセクタ2～26への“Read”コマンドに変換され、HDD2～HDD7ではセクタ4～26への“Read”コマンドに変換され、“Read”コマンドが発行される。次いで処理はステップST24に移る。

ステップST24では、各HDDから読み出されたデータをメモリに格納する。すなわち、読み出し動作が実行され、読み出された全てのデータがメモリ部4に取り込まれる。このとき、スベアセク

タおよび欠陥セクタのデータも一緒にメモリ部4に取り込まれる。  
次いで処理はステップST25に移る。

ステップST25では、メモリに格納されたデータをマッピング表、欠陥リスト、およびスペアリストを参照して読み出す。すなわちメモリ部4内のデータ読出しが行われ、RAID論理アドレスマッピングおよび欠陥/スペアリストと照合して、HDD<sub>1</sub>はセクタ2～26の範囲中で、欠陥セクタ4～5およびスペアセクタ19～21を削除し、2, 3, 6, . . . , 18, 22, 23, . . . , 26の順番に読み出す。また、HDD<sub>2</sub>～HDD<sub>7</sub>は、セクタ4～26の範囲中で、スペアセクタ19～21を削除し、4, 5, 6, . . . , 18, 22, 23, . . . , 26の順番に読み出す。

そして、ステップST25で、その後ユニフィケーションされ、出力される。

次に、データを各HDDに書込む時の動作について説明する。

図14Aに示すようなHDDの20セクタ分のデータを前述のRAID10にセクタアドレスの順に沿って書き込む場合には、HDD<sub>2</sub>～HDD<sub>7</sub>では図12BのRAIDアドレスのうちの区間Aで示される範囲に書き込まれることになり、HDD<sub>1</sub>のディスク上では、図12Cの区間A'で示される範囲に書き込まれることになる。

図14Bは上記の区間Aのセクタアドレスの範囲に書き込まれるデータの並びを表し、図14Cは上記の区間A'のセクタアドレスの範囲に書き込まれるデータの並びを表している。

このとき、RAID10は、図8Aに示すRAID論理アドレスマッピングを参照して、HDD<sub>1</sub>のセクタアドレス2～3(RAIDの割り当てられたアドレスは1～2)、6～18(3～15)、

22～26（16～20）に対して、また、HDD<sub>2</sub>～HDD<sub>7</sub>のセクタアドレス4～18（RAIDの割当られたセクタアドレスは1～15）と、22～26（16～20）に対してデータを書込もうとするが、ブロック内の欠陥セクタおよびブロック間のスペアセクタへの書込みを避けるために、HDD<sub>1</sub>のセクタアドレス2～3、6～18、22～26の3つの領域、およびHDD<sub>2</sub>～HDD<sub>7</sub>のセクタアドレス4～18、22～26の2つの領域に対して、“Write”コマンドが分けて発行されてしまうと、コマンド間で必ず回転待ちが発生し、実行時間が遅れてしまう。

そこで、本発明のデータ記録再生方法では、書込み指定された範囲に欠陥セクタやスペアセクタが含まれる場合には、その欠陥セクタやスペアセクタに対応するデータ部分にダミーデータを配置して連続的に書込みを行うように制御する。図14Bのダミーデータ111および図14Cのダミーデータ112、113がこれに相当する。

このダミーデータ付加処理により、欠陥セクタやスペアセクタを含む範囲が書込み指定された場合にも、含HDDの動作としては、実際には欠陥セクタおよびスペアセクタのスキップ動作を行わないために、HDDの「回転待ち」時間が発生せず、最小限のデータアクセス時間でデータを書き込むことができる。

次に、上述した、データ書込み時の手順を図15を用いて説明する。

図15は、このデータ付加処理の手順を示すフローチャートである。

まずこの処理が開始されるとスラップST11に移る。

ステップST31では、イニシエータからCPU7に“Write”コマンドが発行される。図14A～Cの例ではイニシエータ8から書込み(Write)コマンドが発行される。次いで処理はステップST31に移る。

ステップST32では、発行されたRAIDのセクタアドレスに対応する各HDDのセクタアドレスをRAID理論アドレスマッピング表を参照して検索し、次いでステップST33で検索した各HDDのセクタアドレスへの“Read”コマンドに変換して発行される、図12の例では、RAID10の1～20セクタに対する“Write”コマンドがRAID論理アドレスマッピングと照合される。HDD<sub>1</sub>では、セクタ2～26への“Write”コマンドに変換され、HDD<sub>2</sub>～HDD<sub>7</sub>では、セクタ4～26への“Write”コマンドに変換される。次いで処理はステップST34に移る。

ステップST34では、入力データのストライピング処理が行われ、次いでステップST35でダミーデータを付加してメモリに格納する。すなわち、イニシエータ8から送られた20セクタ分のデータがストライピングされてメモリ部4に送られ、指定区間内のスベアセクタおよび欠陥セクタには、ダミーデータを付加するように制御する。

HDD<sub>1</sub>においては、データが2セクタ送られたら2セクタ分のダミーデータを送り、その後、データが13セクタ送られたら3セクタ分のダミーデータを送り、その後、残りの5セクタのデータをセクタ2～26に書き込むように制御する。

HDD<sub>2</sub>～HDD<sub>7</sub>においては、データが15セクタ送られたら、

3セクタ分のダミーデータを送り、その後、残りの5セクタのデータをメモリ部4に送る。そして、メモリ部4に送られた計23セクタ分のデータを、順番にHDDのセクタ4～26に書き込むように制御する。

そして、ステップST36に移り、メモリ部4内のデータをHDDに書き込んで、処理が終了する。

以上説明した、本発明に係る実施の形態のデータ記録再生方法およびデータ装置データ書込み／読み出し手順によれば、RAIDに適用して、その動作を停止することなく欠陥処理を行うことができ、しかも欠陥処理後にディスクアクセスが遅くなることがない。

従って、欠陥処理後のヘッドのシーク動作及び回転待ち動作によるディスクへのアクセス劣化が発生せず、しかもRAIDが通常動作を行いながら欠陥処理の割り込みが可能となる。すなわち、マルチチャンネルの使用においては、そのうちの1チャンネル分だけを使って欠陥処理を行うことができ、残りのチャンネルの運用を止める必要がなくなる。

さらに、1チャンネル分を欠陥処理専用として使用可能な場合には、イニシエータ8が自動で欠陥処理コマンドを、この専用チャンネルに発行するように制御すれば、欠陥処理が自動で実施され、ユーザは全く欠陥を意識することなく運用継続できる。

なお、本発明は上記実施の形態のみに限定されるものではなく、本発明の要旨を逸脱しない範囲において種々の変更が可能であることは勿論である。

### 請 求 の 範 囲

1. 複数のディスクから構成されたディスクドライブの各ディスクのトラックが複数のブロックに分割され、上記ブロックの間には所定のセクタ数からなるスペアセクタが所定の間隔で予め配置され、上記ディスクドライブにデータを分配して記録及び／又は再生するデータ記録再生方法であって、

上記ディスクドライブの欠陥セクタを検出する欠陥セクタ検出工程と、

検出された欠陥セクタのアドレスを欠陥リストに登録する欠陥セクタ登録工程と、

上記欠陥セクタを含むブロックを上記欠陥リストと予め設定された対応関係とに基づいて、欠陥セクタを除いてディスク上で連続して配置されるよう上記スペアセクタを用いて代替領域に再登録するリアサイン工程と

を有することを特徴とするデータ記録再生方法。

2. 上記再登録されたブロックに含まれている欠陥セクタのデータを復元する復元工程と、

上記復元データを上記代替領域に再登録されたブロックに再記録するリビルト工程と

をさらに有することを特徴とする請求の範囲第1項記載のデータ記録再生方法。

3. 上記リアサイン工程では、上記欠陥セクタ数が上記スペアセクタ数よりも多いときには、上記欠陥セクタを含むブロックの再登録



を複数のスペアセクタを用いて行うことを特徴とする請求の範囲第1項記載のデータ記録再生方法。

4. 上記代替領域に再登録された復元データを再生するときには、  
読出し指定された範囲内の、スペアセクタ及び欠陥セクタを含む全セクタからデータを連続的に読出してメモリに格納し、メモリに格納されたデータを読み出すときに、上記欠陥セクタのデータを削除することを特徴とする請求の範囲第2項記載のデータ記録再生方法。

5. 上記欠陥セクタを含む領域にデータを記録するときには、書込み指定された範囲内の、スペアセクタ及び欠陥セクタを含む全セクタにデータを連続的に書込み、上記欠陥セクタにはダミーデータを書き込むことを特徴とする請求の範囲第2項記載のデータ記録再生方法。

6. 複数枚のディスクから構成されたディスクドライブの各ディスクのトラックが複数のブロックに分割され、上記ブロックの間には所定のセクタ数からなるスペアセクタが所定の間隔で予め配置され、上記ディスクドライブにデータを分配して記録及び／又は再生するデータ記録再生装置において、

上記ディスクドライブの論理セクタと物理セクタとの対応を記憶するアドレスマッピング手段と、

上記ディスクドライブの欠陥セクタを検出して、該欠陥セクタのアドレスを記憶する欠陥セクタ記憶手段と、

上記欠陥セクタを含むブロックを上記アドレスマッピング手段と上記欠陥セクタ記憶手段とに基づいて、上記ディスク上で上記欠陥セクタを除いて連続して配置されるように上記スペアセクタを用い

て代替領域に再登録するリアサイン手段と

を有することを特徴とするデータ記録再生装置。

7. 上記複数のディスクドライブの記録／再生データを記憶する記憶手段をさらに備え、

上記代替領域に再記録された復元データを再生するときには、読出し指定された範囲内の、スベアセクタ及び欠陥セクタを含む全セクタからデータを連続的に読出して上記記憶手段に格納し、格納されたデータを読み出すときに上記欠陥セクタのデータを削除することを特徴とする請求の範囲第6項記載のデータ記録再生装置。

8. 上記欠陥セクタを含む領域にデータを記録するときには、書込み指定された範囲内の、スベアセクタ及び欠陥セクタを含む全セクタにデータを連続的に書込み、上記欠陥セクタにはダミーデータを書き込むことを特徴とする請求の範囲6記載のデータ記録再生装置。

9. 上記再登録されたブロックに含まれている欠陥セクタのデータを復元する復元手段と、

上記復元されたデータを、上記代替領域に再登録されたブロックに再記録するリビルト手段と

をさらに有することを特徴とする請求の範囲第6項記載のデータ記録再生装置。

10. 上記複数のディスクドライブの記録／再生データを記憶する記憶手段をさらに備え、

上記代替領域に再記録された復元データを再生するときには、読出し指定された範囲内の、スベアセクタ及び欠陥セクタを含む全セクタからデータを連続的に読出して上記記憶手段に格納し、格納されたデータを読み出すときに上記欠陥セクタのデータを削除し、上

記欠陥セクタを含む領域にデータを記録するときには、書込み指定された範囲内の、スベアセクタ及び欠陥セクタを含む全セクタにデータを連続的に書込み、上記欠陥セクタにはダミーデータを書き込むことを特徴とする請求の範囲第9項記載のデータ記録再生装置。

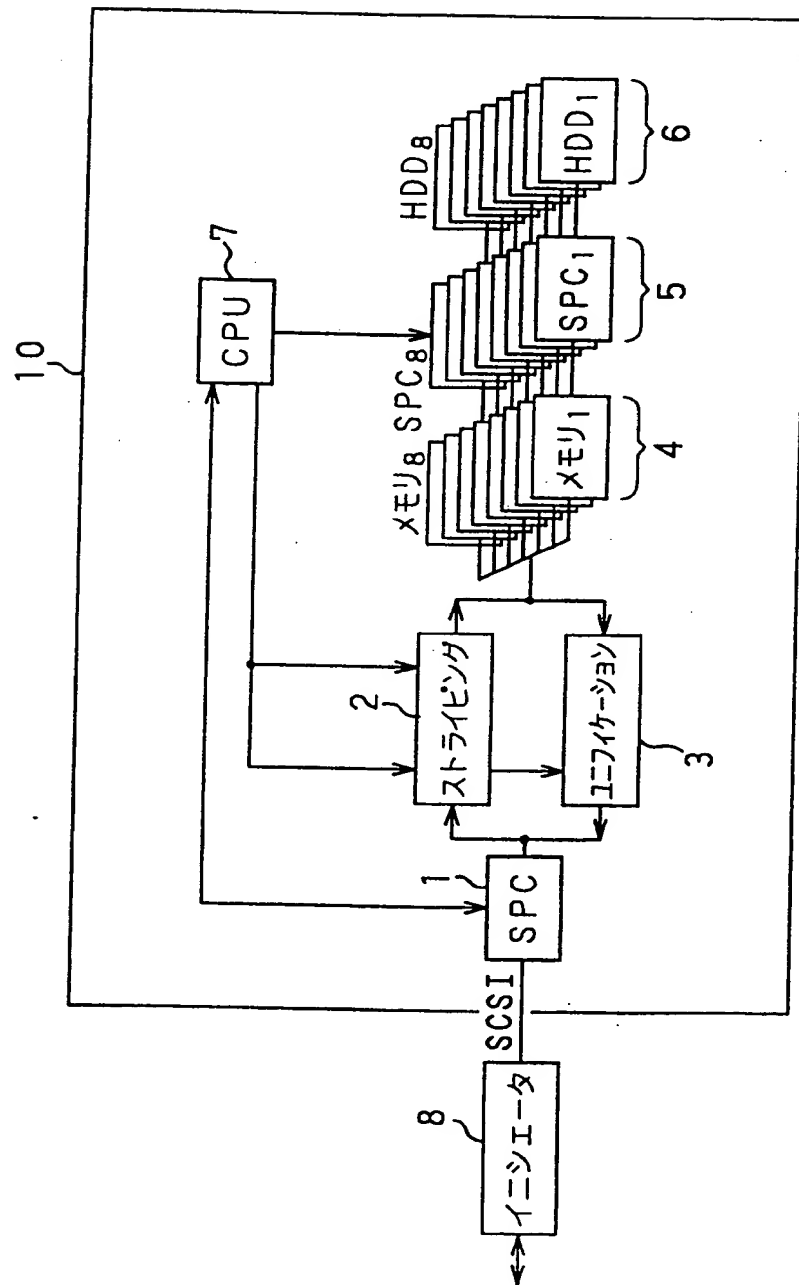


FIG. 1

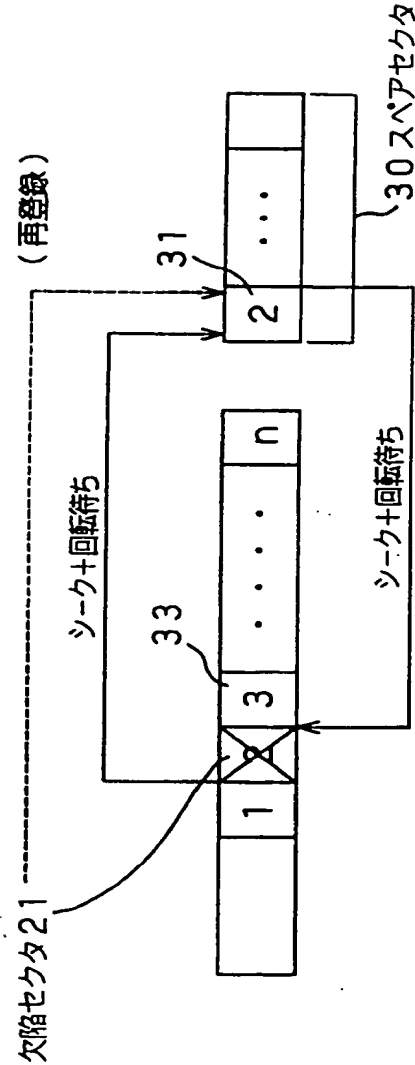


FIG. 2

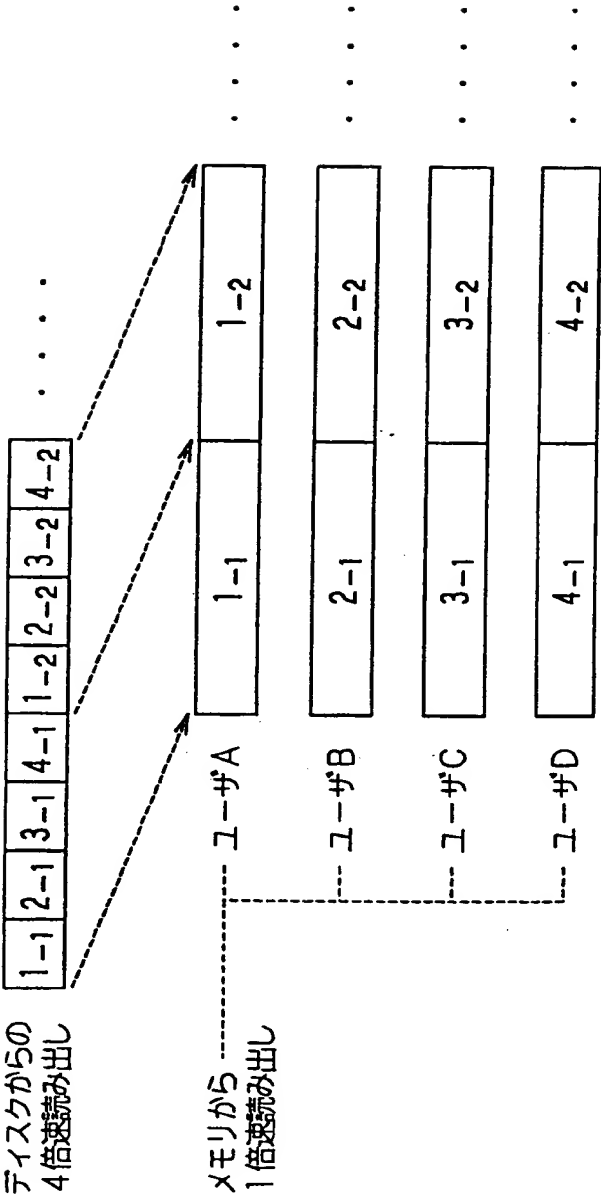


FIG. 3

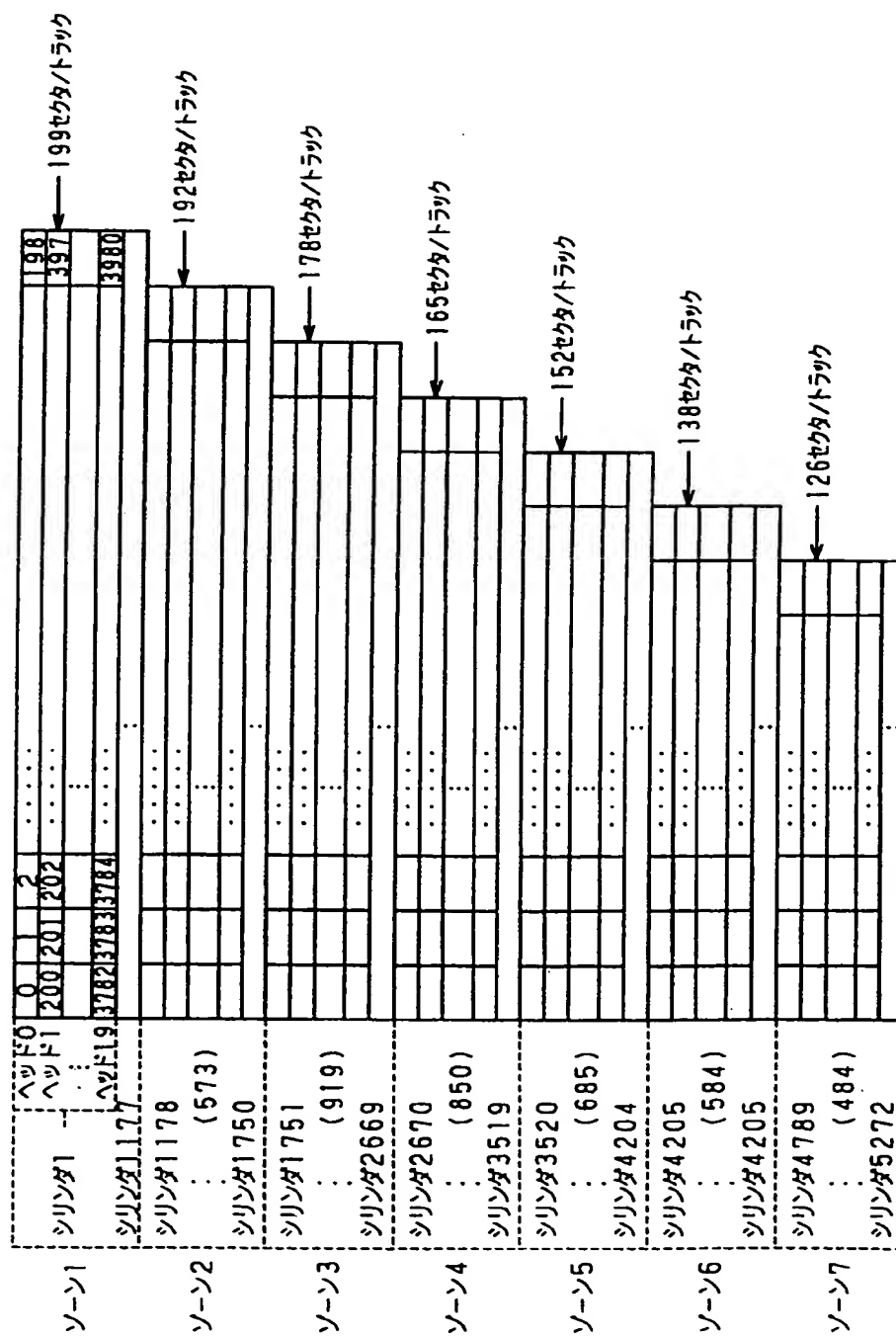


FIG. 4

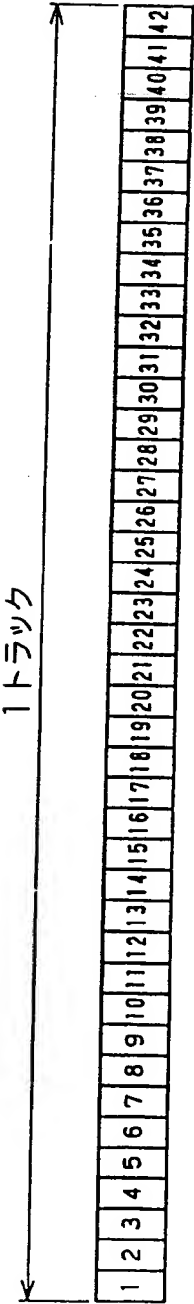


FIG. 5A

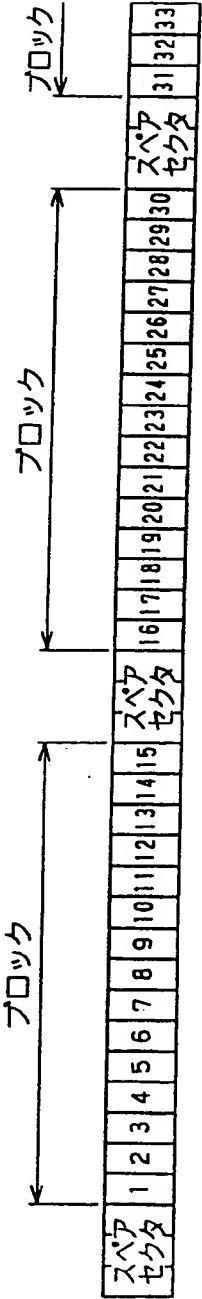


FIG. 5B



RAID論理アドレスマッピング	
RAIDセクタ アドレス	各HDDのセクタアドレス
1	1 2 3 4 5 6 7 8 9 10 11 12 13 14 15 16 17 18 19 20 21 22 23 24 25 26 27 28 29 30 31 32 33 34 35 36 37 38 39 40 41 42 43 44 45 46 47 48 49 50 51 52 53 54 55 56 57 58 59 60 61 62 63 64 65 66 67 68 69 70 71 72 73 74 75 76 77 78 79 80 81 82 83 84 85 86 87 88 89 90 91 92 93 94 95 96 97 98 99 100
2	101 102 103 104 105 106 107 108 109 110 111 112 113 114 115 116 117 118 119 120 121 122 123 124 125 126 127 128 129 130 131 132 133 134 135 136 137 138 139 140 141 142 143 144 145 146 147 148 149 150 151 152 153 154 155 156 157 158 159 160 161 162 163 164 165 166 167 168 169 170 171 172 173 174 175 176 177 178 179 180 181 182 183 184 185 186 187 188 189 190 191 192 193 194 195 196 197 198 199 200
3	201 202 203 204 205 206 207 208 209 210 211 212 213 214 215 216 217 218 219 220 221 222 223 224 225 226 227 228 229 230 231 232 233 234 235 236 237 238 239 240 241 242 243 244 245 246 247 248 249 250 251 252 253 254 255 256 257 258 259 260 261 262 263 264 265 266 267 268 269 270 271 272 273 274 275 276 277 278 279 280 281 282 283 284 285 286 287 288 289 290 291 292 293 294 295 296 297 298 299 300
4	301 302 303 304 305 306 307 308 309 310 311 312 313 314 315 316 317 318 319 320 321 322 323 324 325 326 327 328 329 330 331 332 333 334 335 336 337 338 339 340 341 342 343 344 345 346 347 348 349 350 351 352 353 354 355 356 357 358 359 360 361 362 363 364 365 366 367 368 369 370 371 372 373 374 375 376 377 378 379 380 381 382 383 384 385 386 387 388 389 390 391 392 393 394 395 396 397 398 399 400
:	:
16	401 402 403 404 405 406 407 408 409 410 411 412 413 414 415 416 417 418 419 420 421 422 423 424 425 426 427 428 429 430 431 432 433 434 435 436 437 438 439 440 441 442 443 444 445 446 447 448 449 450 451 452 453 454 455 456 457 458 459 460 461 462 463 464 465 466 467 468 469 470 471 472 473 474 475 476 477 478 479 480 481 482 483 484 485 486 487 488 489 490 491 492 493 494 495 496 497 498 499 500
17	501 502 503 504 505 506 507 508 509 510 511 512 513 514 515 516 517 518 519 520 521 522 523 524 525 526 527 528 529 530 531 532 533 534 535 536 537 538 539 540 541 542 543 544 545 546 547 548 549 550 551 552 553 554 555 556 557 558 559 560 561 562 563 564 565 566 567 568 569 570 571 572 573 574 575 576 577 578 579 580 581 582 583 584 585 586 587 588 589 590 591 592 593 594 595 596 597 598 599 600
:	:
30	601 602 603 604 605 606 607 608 609 610 611 612 613 614 615 616 617 618 619 620 621 622 623 624 625 626 627 628 629 630 631 632 633 634 635 636 637 638 639 640 641 642 643 644 645 646 647 648 649 650 651 652 653 654 655 656 657 658 659 660 661 662 663 664 665 666 667 668 669 670 671 672 673 674 675 676 677 678 679 680 681 682 683 684 685 686 687 688 689 690 691 692 693 694 695 696 697 698 699 700
:	:

FIG. 6A

HDD 欠陥リスト
HDD1
HDD2
HDD3
HDD4
HDD5
HDD6
HDD7
HDD8

FIG. 6B

HDD スベアリスト
HDD1 1,2,3...
HDD2 1,2,3...
HDD3 1,2,3...
HDD4 1,2,3...
HDD5 1,2,3...
HDD6 1,2,3...
HDD7 1,2,3...
HDD8 1,2,3...

FIG. 6C

RAID 欠陥リスト

FIG. 6D

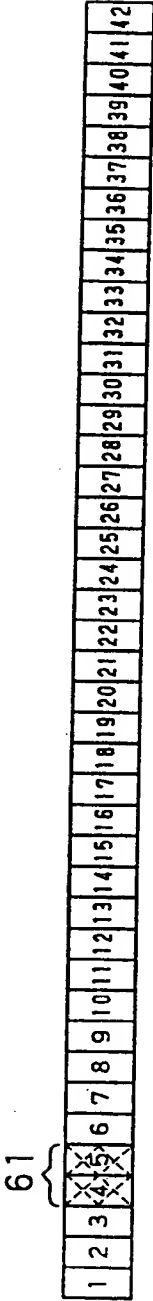


FIG. 7A

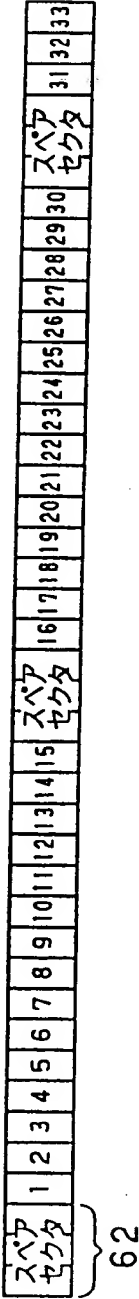


FIG. 7B

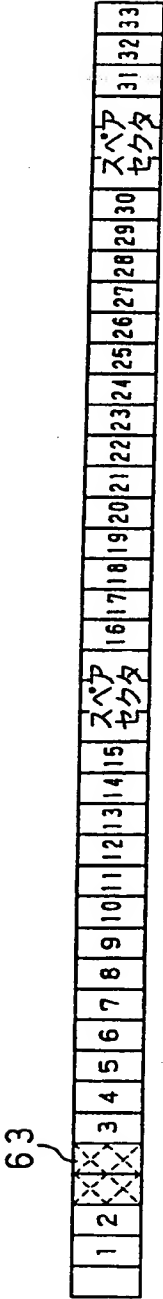


FIG. 7C

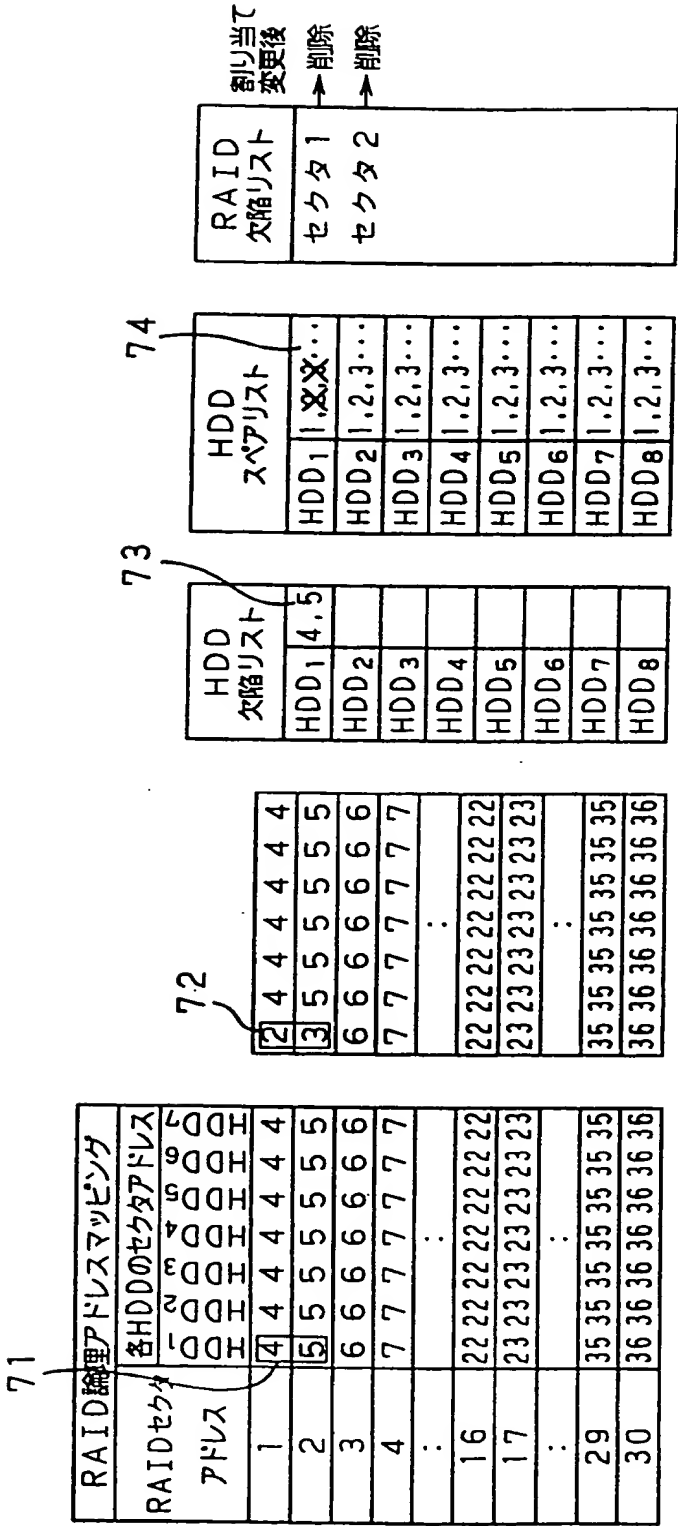


FIG. 8A      FIG. 8B      FIG. 8C      FIG. 8D      FIG. 8E

9/15

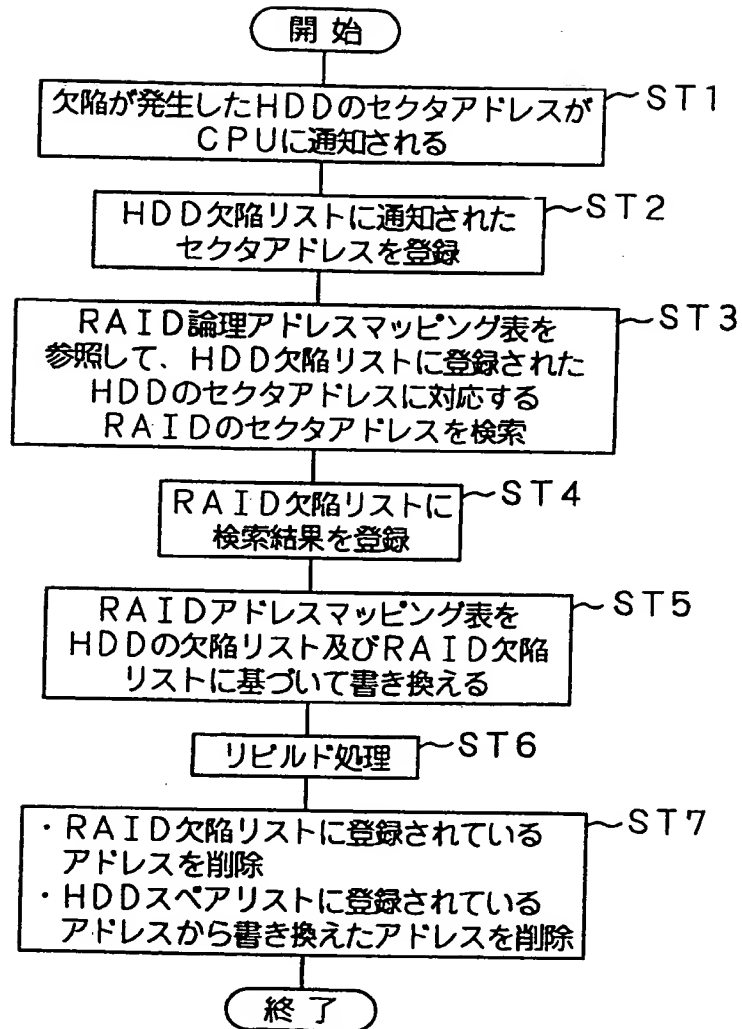


FIG. 9

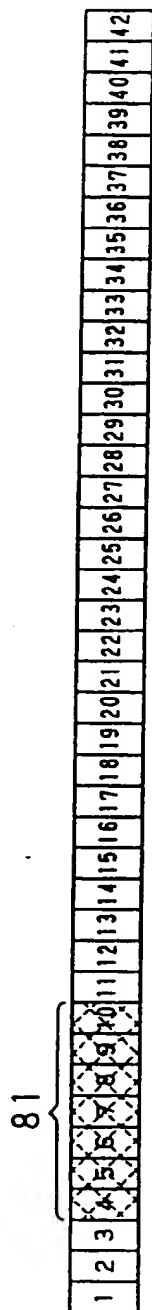


FIG. 10A

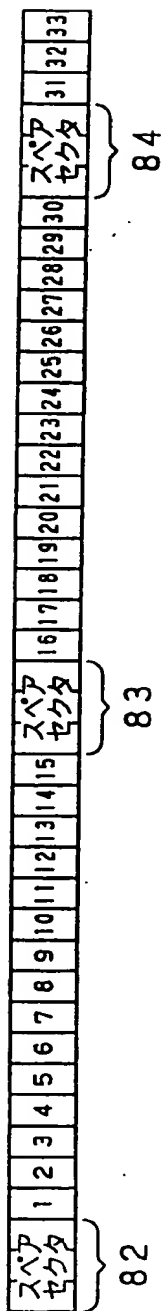


FIG. 10B

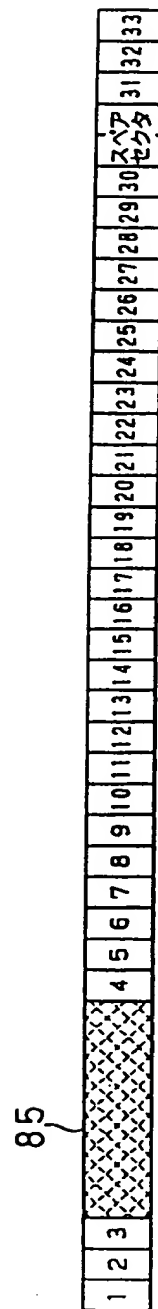


FIG. 10C

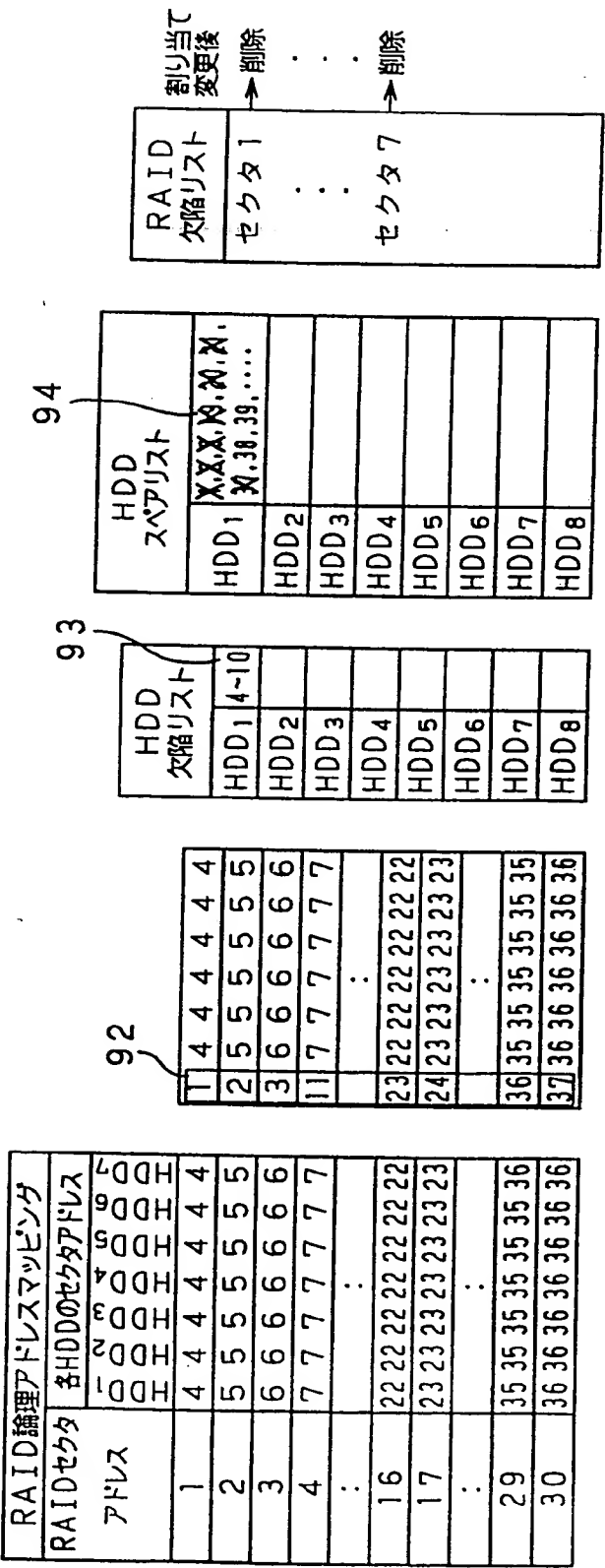


FIG. 11A      FIG. 11B      FIG. 11C      FIG. 11D      FIG. 11E

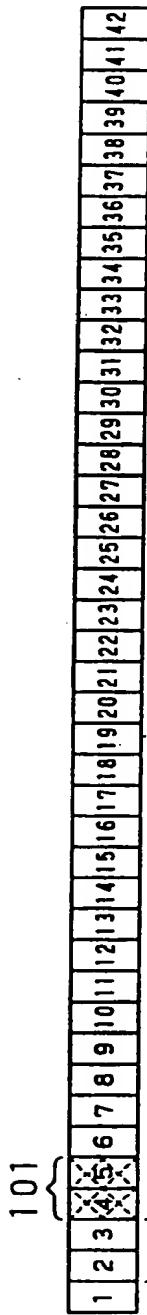


FIG. 12A

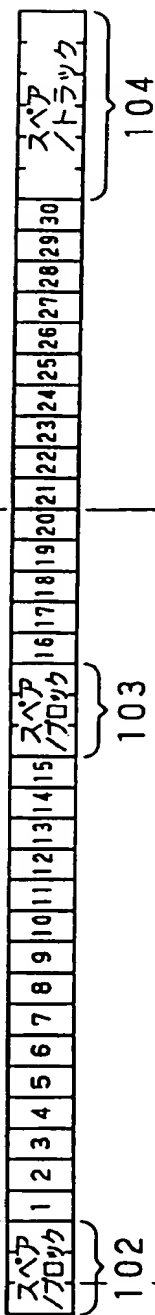


FIG. 12B

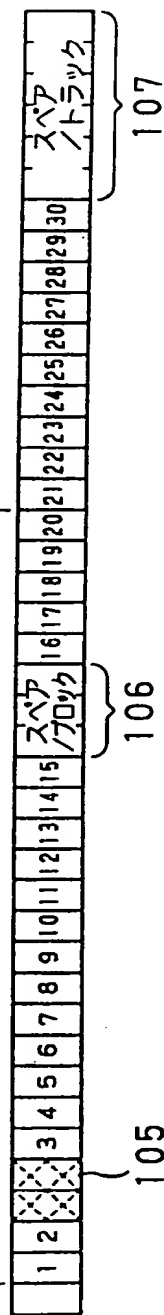


FIG. 12C

13/15

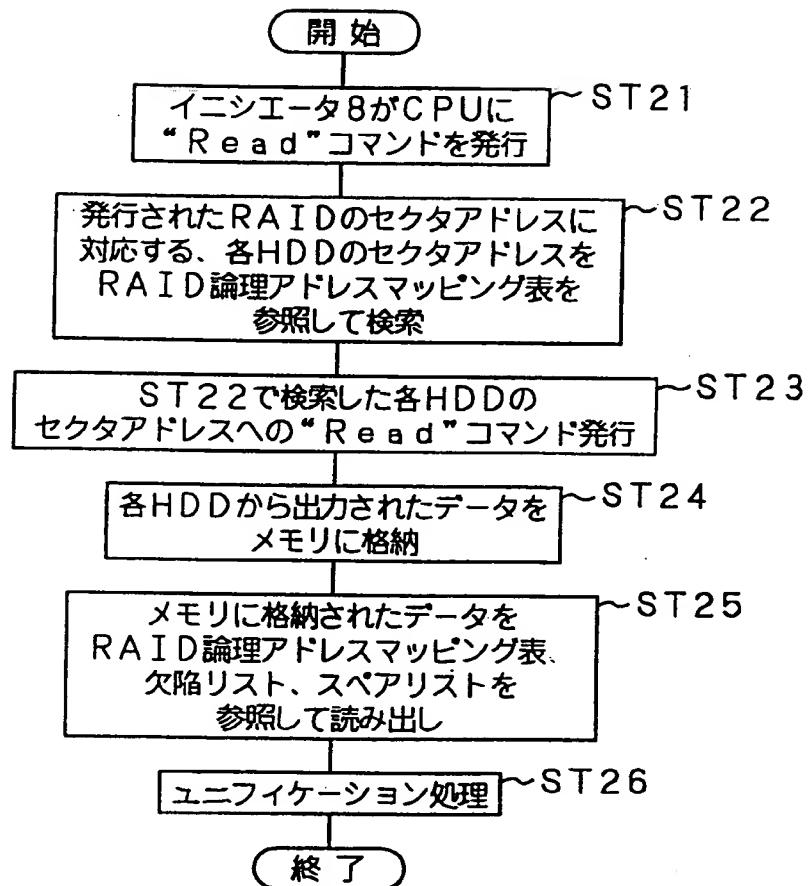


FIG. 13



FIG. 14A

1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20
---	---	---	---	---	---	---	---	---	----	----	----	----	----	----	----	----	----	----	----

FIG. 14B

1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20
---	---	---	---	---	---	---	---	---	----	----	----	----	----	----	----	----	----	----	----

FIG. 14C

1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20
---	---	---	---	---	---	---	---	---	----	----	----	----	----	----	----	----	----	----	----

15/15

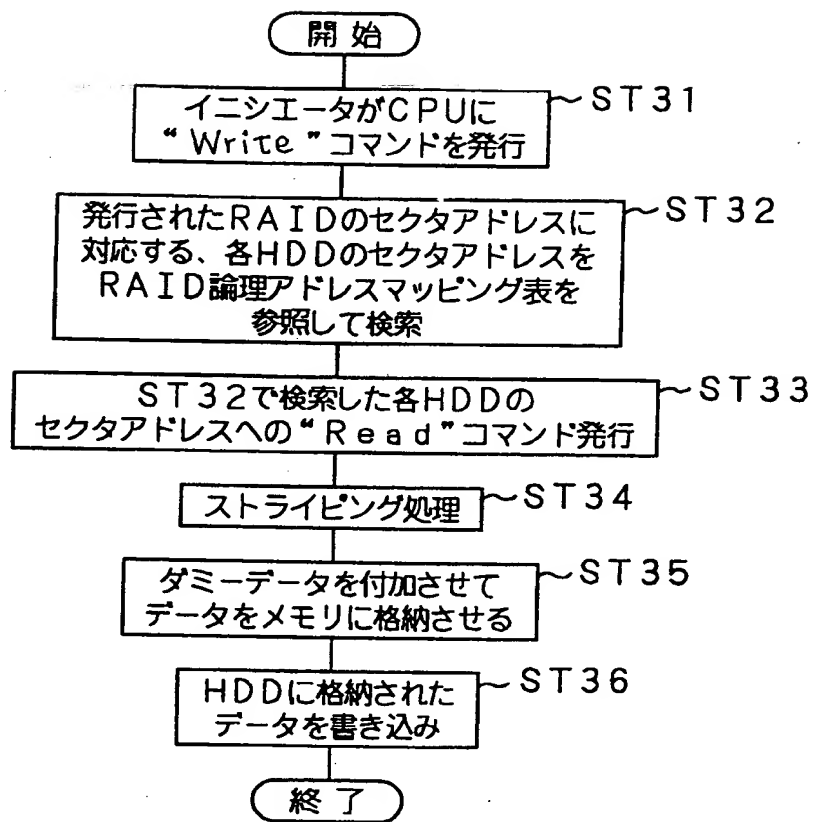


FIG. 15

## INTERNATIONAL SEARCH REPORT

International application No.

PCT/JP98/00627

A. CLASSIFICATION OF SUBJECT MATTER  
Int.Cl<sup>6</sup> G11B20/12, G11B20/10

According to International Patent Classification (IPC) or to both national classification and IPC

## B. FIELDS SEARCHED

Minimum documentation searched (classification system followed by classification symbols)

Int.Cl<sup>6</sup> G11B20/12, G11B20/10, G11B20/18, G06F3/06, G06F12/16

Documentation searched other than minimum documentation to the extent that such documents are included in the fields searched

Jitsuyo Shinan Koho 1926-1998 Toroku Jitsuyo Shinan Koho 1994-1998

Kokai Jitsuyo Shinan Koho 1971-1998

Electronic data base consulted during the international search (name of data base and, where practicable, search terms used)

## C. DOCUMENTS CONSIDERED TO BE RELEVANT

Category*	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
Y	JP, 6-110619, A (Fujitsu Ltd.), April 22, 1994 (22. 04. 94), Column 8, line 2 to column 10, line 40 (Family: none)	1-3, 6, 9
A		4, 5, 7, 8, 10
Y	JP, 5-334017, A (Fujitsu Ltd.), December 17, 1993 (17. 12. 93), Column 4, lines 2 to 50 (Family: none)	1-3, 6, 9
A		4, 5, 7, 8, 10
Y	JP, 6-76480, A (Fujitsu Ltd.), March 18, 1994 (18. 03. 94), Column 6, line 9 to column 8, line 23 (Family: none)	1-3, 6, 9
A		4, 5, 7, 8, 10

☒ Further documents are listed in the continuation of Box C.☐ See patent family annex.

\* Special categories of cited documents:

- "A" document defining the general state of the art which is not considered to be of particular relevance
- "E" earlier document but published on or after the international filing date
- "L" document which may throw doubts on priority claim(s) or which is cited to establish the publication date of another citation or other special reason (as specified)
- "O" document referring to an oral disclosure, use, exhibition or other means
- "P" document published prior to the international filing date but later than the priority date claimed

- "T" later document published after the international filing date or priority date and not in conflict with the application but cited to understand the principle or theory underlying the invention
- "X" document of particular relevance; the claimed invention cannot be considered novel or cannot be considered to involve an inventive step when the document is taken alone
- "Y" document of particular relevance; the claimed invention cannot be considered to involve an inventive step when the document is combined with one or more other such documents, such combination being obvious to a person skilled in the art
- "&" document member of the same patent family

Date of the actual completion of the international search  
May 12, 1998 (12. 05. 98)Date of mailing of the international search report  
May 26, 1998 (26. 05. 98)Name and mailing address of the ISA/  
Japanese Patent Office

Authorized officer

Facsimile No.

Telephone No.

## INTERNATIONAL SEARCH REPORT

International application No.

PCT/JP98/00627

C (Continuation). DOCUMENTS CONSIDERED TO BE RELEVANT

Category*	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
Y	JP, 6-111479, A (Nikon Corp.), April 22, 1994 (22. 04. 94), Column 7, line 4 to column 8, line 6 (Family: none)	1-3, 6, 9
A		4, 5, 7, 8, 10
A	JP, 5-28646, A (Fujitsu Ltd.), February 5, 1993 (05. 02. 93) (Family: none)	1-10
A	JP, 2-218068, A (NEC Hokkaido Software K.K.), August 30, 1990 (30. 08. 90) (Family: none)	1-10
A	JP, 62-46468, A (Hitachi, Ltd.), February 28, 1987 (28. 02. 87) (Family: none)	1-10
A	JP, 61-214280, A (Siemens AG.), September 24, 1986 (24. 09. 86) & EP, 195324, A & US, 4706136, A	1-10
A	JP, 58-3104, A (Toshiba Corp.), January 8, 1983 (08. 01. 83) (Family: none)	1-10

## A. 発明の属する分野の分類 (国際特許分類 (IPC))

Int Cl<sup>8</sup> G11B20/12, G11B20/10

## B. 調査を行った分野

調査を行った最小限資料 (国際特許分類 (IPC))

Int Cl<sup>8</sup> G11B20/12, G11B20/10, G11B20/18,  
G06F3/06, G06F12/16

最小限資料以外の資料で調査を行った分野に含まれるもの

日本国実用新案公報 1926-1998年  
日本国公開実用新案公報 1971-1998年  
日本国登録実用新案公報 1994-1998年

国際調査で使用した電子データベース (データベースの名称、調査に使用した用語)

## C. 関連すると認められる文献

引用文献の カテゴリー*	引用文献名 及び一部の箇所が関連するときは、その関連する箇所の表示	関連する 請求の範囲の番号
Y	J P, 6-110619, A (富士通株式会社) 22. 4月. 1994 (22. 04. 94) 第8欄第2行-第10欄第40行 (ファミリーなし)	1-3, 6, 9
A		4, 5, 7, 8, 10

☒ C欄の続きにも文献が列挙されている。☐ パテントファミリーに関する別紙を参照。

## \* 引用文献のカテゴリー

- 「A」 特に関連のある文献ではなく、一般的技術水準を示すもの  
「E」 先行文献ではあるが、国際出願日以後に公表されたもの  
「L」 優先権主張に疑義を提起する文献又は他の文献の発行日若しくは他の特別な理由を確立するために引用する文献 (理由を付す)  
「O」 口頭による開示、使用、展示等に言及する文献  
「P」 国際出願日前で、かつ優先権の主張の基礎となる出願

- の日の後に公表された文献  
「T」 国際出願日又は優先日後に公表された文献であって出願と矛盾するものではなく、発明の原理又は理論の理解のために引用するもの  
「X」 特に関連のある文献であって、当該文献のみで発明の新規性又は進歩性がないと考えられるもの  
「Y」 特に関連のある文献であって、当該文献と他の1以上の文献との、当業者にとって自明である組合せによって進歩性がないと考えられるもの  
「&」 同一パテントファミリー文献

国際調査を完了した日

12. 05. 98

国際調査報告の発送日

26.05.98

国際調査機関の名称及びあて先

日本国特許庁 (ISA/J P)  
郵便番号100-8915  
東京都千代田区霞が関三丁目4番3号

特許庁審査官 (権限のある職員)

早川 卓哉

5D

9295

電話番号 03-3581-1101 内線 3552

C (続き) . 関連すると認められる文献		
引用文献の カテゴリー*	引用文献名 及び一部の箇所が関連するときは、その関連する箇所の表示	関連する 請求の範囲の番号
Y A	JP, 5-334017, A (富士通株式会社) 17. 12月. 1993 (17. 12. 93) 第4欄第2行-第50行 (ファミリーなし)	1-3, 6, 9
Y A	JP, 6-76480, A (富士通株式会社) 18. 3月. 1994 (18. 03. 94) 第6欄第9行-第8欄第23行 (ファミリーなし)	4, 5, 7, 8, 10
Y A	JP, 6-111479, A (株式会社ニコン) 22. 4月. 1994 (22. 04. 94) 第7欄第4行-第8欄第6行 (ファミリーなし)	1-3, 6, 9
A	JP, 5-28646, A (富士通株式会社) 5. 2月. 1993 (05. 02. 93) (ファミリーなし)	4, 5, 7, 8, 10
A	JP, 2-218068, A (北海道日本電気ソフトウェア株式会社) 30. 8月. 1990 (30. 08. 90) (ファミリーなし)	1-10
A	JP, 62-46468, A (株式会社日立製作所) 28. 2月. 1987 (28. 02. 87) (ファミリーなし)	1-10
A	JP, 61-214280, A (シーメンス、アクチエンゲゼルシャフト) 24. 9月. 1986 (24. 09. 86) & EP, 195324, A&US, 4706136, A	1-10
A	JP, 58-3104, A (株式会社東芝) 8. 1月. 1983 (08. 01. 83) (ファミリーなし)	1-10